

HTTP parametrizacija

HTTP parametri i njihova sintaksa su jasno definisani u dokumentaciji samog protokola, pa tako postoji poseban set pravila koji parametri moraju ispoštovati kako bi se uspešno prosledili. Ovi parametri koji se šalju u zaglavlju odgovora su nam bitni kako bi se standardizovao protok i format informacija prilikom zahteva/odgovora.

HTTP verzija

Sintaksa za prosleđivanje HTTP verzije je takva da se prosleđuje prvo verzija protokola, a onda i podverzija (major i minor) protokola odvojena tačkom (elementi u navodnicima su delovi sintakse; razmaci između elemenata pod navodnicima nisu deo sintakse, već su tu radi lakšeg čitanja):

Sintaksa vrednosti HTTP verzije:

```
"HTTP" "/" 1* "." 1*
```

Primer:

```
HTTP/2.4  
HTTP/2.13
```

U ovom primeru se vidi da se brojevi verzije i podverzije tretiraju kao posebni brojevi i ti brojevi se mogu uvećavati tako da umesto jedne cifre imamo dve (što se vidi i na primeru). Takođe, po HTTP standardu, vodeće nule u brojevima verzija (01, 02...) se ne šalju.

Aplikacija koja šalje zahtev ili odgovor koja uključuje vrednost zaglavlja HTTP-Version, kao na primer HTTP/1.1, mora biti dizajnirana tako da može podržavati maksimalno tu verziju.

Jedinstveni identifikator resursa – URI (Uniform Resource Identifier)

Jedinstveni identifikator resursa identifikuje stranicu po lokaciji, imenu ili po oba kriterijuma zajedno, pa se tako može reći da se URI sastoji od URL-a (jedinstvenog identifikatora lokacije) i URN-a (jedinstvenog identifikatora imena). URI-ovi se po HTTP standardu mogu prikazati kao apsolutne ili relativne putanje. Ove dve forme se razlikuju po tome što apsolutne URI putanje uvek započinju imenom šeme (http, https, ftp), nakon kojeg slede dve tačke. Sam HTTP protokol ne nameće limit za dužinu URI-a, te serveri moraju biti dizajnirani za rad sa takvim scenarijima. Takođe je na serverima i da budu u stanju da vrate grešku 414 – URI Too Long ako je dužina URI-a veća nego ona za čiji rad su serveri projektovani da rade.

Šema http ili https nam govori da je reč o HTTP protokolu. Sintaksa URL-a izgleda ovako (elementi u uglastim zagrada su opcioni):

```
"http:" "://" host [ ":" broj porta ] [ apsolutna_putanja [ "?" upit ] ]
```

Po specifikaciji HTTP protokola, ako broj porta nije prosleđen, koristi se podrazumevani port 80. Iako je moguće koristiti IP adresu kao host, praksa je da se takvi scenariji izbegavaju kad god je to moguće. Ako je apsolutna putanja izostavljena, mora se zameniti znakom / u samom URL-u. Upiti se prosleđuju po pravilima po kojima funkcionišu upitni stringovi, o kojima je već bilo reči u kursu *Python Net Programming*.

Datum i vreme

HTTP standard podržava tri različita formata datuma:

- Sun, 26 Jul 2020 08:49:37 GMT
- Sunday, 26-Jul-20 08:49:37 GMT
- Sun Jul 26 08:49:37 2020

Prvi format se preferira kao svojevrsni internet standard i predstavlja format fiksirane dužine. Drugi format se takođe dosta koristi, ali mu je mana to što se godine prosleđuju kao dvocifreni brojevi. Klijenti i serveri bazirani na HTTP/1.1 verziji protokola moraju podržavati pravilno raščlanjivanje sva tri formata zbog kompatibilnosti sa HTTP/1.0, ali moraju generisati samo prvi format kada je reč o datumu i vremenu u zaglavlju. Takođe, sve vremenske oznake u HTTP-u moraju biti predstavljene u griničkom vremenu. Imena dana u nedelji, kao i imena meseci, osetljiva su na upotrebu malih/velikih slova.

Sintaksa za datume preuzeta direktno iz protokolne dokumentacije izgleda ovako:

Sintaksa vrednosti Content-Type polje:

```
HTTP-date      = rfc1123-date | rfc850-date | asctime-date
rfc1123-date   = wkday "," SP date1 SP time SP "GMT"
rfc850-date    = weekday "," SP date2 SP time SP "GMT"
asctime-date   = wkday SP date3 SP time SP 4DIGIT
date1          = 2DIGIT SP month SP 4DIGIT
                ; day month year (e.g., 02 Jun 1982)
date2          = 2DIGIT "-" month "-" 2DIGIT
                ; day-month-year (e.g., 02-Jun-82)
date3          = month SP ( 2DIGIT | ( SP 1DIGIT ) )
                ; month day (e.g., Jun 2)
time           = 2DIGIT ":" 2DIGIT ":" 2DIGIT
                ; 00:00:00 - 23:59:59
wkday          = "Mon" | "Tue" | "Wed"
                | "Thu" | "Fri" | "Sat" | "Sun"
weekday        = "Monday" | "Tuesday" | "Wednesday"
                | "Thursday" | "Friday" | "Saturday" | "Sunday"
month          = "Jan" | "Feb" | "Mar" | "Apr"
                | "May" | "Jun" | "Jul" | "Aug"
                | "Sep" | "Oct" | "Nov" | "Dec"
```

Primer:

```
Sun, 10 Oct 2021 02:03:00 GMT = wkday "," SP date1 SP time SP "GMT"
Sunday, 10-Oct-21 02:03:00 GMT = weekday "," SP date2 SP time SP
"GMT"
Sun Oct 10 02:03:00 2021 = wkday SP date3 SP time SP 4DIGIT
```

Pitanje

URI ne sadrži (izabrati jedan odgovor):

- URN
- **RFC**
- URL

Objašnjenje:

Tačan odgovor je da URI ne sadrži iz RFC (RFC predstavlja publikaciju Internet društva (ISOC)).

Skupovi znakova (character sets)

Ovim HTTP parametrom određujemo koji skup znakova klijent/server podržava. Informacija o skupu znakova se šalje preko Accept-Charset polja zaglavlja. Više ovih setova se može proslediti po parametru, ali se onda moraju odvojiti zapetom. Ako ovo polje nije prosleđeno, podrazumeva se da je skup znakova US-ASCII. Lista skupova znakova i znakova koji im pripadaju se može naći na ovom [linku](#). Neki od postojećih skupova znakova su:

- US-ASCII
- ISO-8859-1
- ISO-8859-7

Kodiranje sadržaja (content encoding)

Ovo polje nam pokazuje koji algoritam kodiranja se koristio za kodiranje sadržaja pre nego što je sadržaj poslat. Svrha ovog procesa je kompresija čitavog sadržaja (dokumenta) koji se šalje, a da se pri tome ne izgube originalni podaci. Ovako kodirani podaci se šalju u poljima zaglavlja: Accept-Encoding i Content-Encoding, pa tako ovo polje pokazuje ime algoritma koji je korišćen u kompresiji koji se može upotrebiti za dekodovanje i prebacivanje dokumenta u originalnu formu. Slede neki od češće korišćenih algoritama kompresije prema organizaciji IANA (Internet Assigned Numbers Authority):

- gzip – format kompresije koji je razvijen za GNU zip; baziran je na Lempel-Ziv algoritmu (LZ77);
- compress – format kompresije baziran na Lempel-Ziv-Welch (LZW) algoritmu;
- deflate – format kompresije baziran na istoimenom algoritmu kompresije;
- identity – prisustvo ove vrednosti ukazuje na to da kompresija nije primenjena.

Primer vrednosti Accept-Encoding polje:

```
Accept-Encoding: gzip  
Accept-Encoding: deflate, gzip
```

Tipovi sadržaja

Na tipove sadržaja odnosi se polje u zaglavlju koje nam ukazuje na to u kom formatu je sadržaj koji se šalje. Content-Type je polje zaglavlja u kojem se šalju informacije o tipu sadržaja. Sintaksa za vrednost ovog polja izgleda ovako (element nakon zvezdice je opcioni):

Sintaksa vrednosti Content-Type polja:

```
tip "/" podtip *( ";" parametar )
```

Primer:

```
Content-Type: text/richtext; charset=UTF-8
```

Svi elementi u sintaksi su osetljivi na velika/mala slova, osim vrednosti koja ide umesto parametar elementa. Lista imena tipova i podtipova koji se mogu slati se može naći na [linku](#).

Oznake jezika

HTTP protokol nam omogućava da u zaglavlju pošaljemo i informaciju koja pokazuje na kom jeziku ili za koji jezik je sadržaj kreiran. Takođe, ako je reč o zahtevu – oznakom jezika šaljemo informaciju serveru na kom jeziku želimo povratnu informaciju i slično. Oznake jezika se šalju u Accept-Language ili Content-Language poljima zaglavlja. Sintaksa po kojoj se ovom zaglavlju prosleđuju vrednosti je sledeća (element nakon zvezdice je opcioni):

Sintaksa vrednosti za Accept-Language polje:

```
glavna_oznaka *( "-" pod_oznaka )
```

Primer:

```
Accept-Language: en  
Accept-Language: en-US
```

I ovi elementi su osetljivi na mala/velika slova, pa o tome treba posebno voditi računa. Lista dostupnih oznaka i podoznaka jezika koju je definisala IANA nalazi se na ovom [linku](#).

Više detalja o HTTP parametrima i zaglavljima može se pronaći na [ovoj stranici](#).

Rezime

- HTTP parametri i njihova sintaksa su jasno definisani u dokumentaciji samog protokola.
- Sintaksa za prosleđivanje HTTP verzije je takva da se prosleđuje prvo verzija protokola, a onda i podverzija (major i minor) protokola, odvojena tačkom.
- Jedinstveni identifikator resursa identifikuje stranicu po lokaciji, imenu ili po oba kriterijuma zajedno.

- URI se sastoji od URL-a (jedinственог идентификатора lokacije) i URN-a (jedinственог идентификатора imena).
- HTTP standard podržava tri različita formata datuma:
 - Sun, 26 Jul 2020 08:49:37 GMT
 - Sunday, 26-Jul-20 08:49:37 GMT
 - Sun Jul 26 08:49:37 2020
- URI-ovi se po HTTP standardu mogu prikazati kao apsolutne ili relativne putanje.
- Informacija o skupu znakova se šalje preko Accept-Charset polja zaglavlja.
- Kodiranje sadržaja nam pokazuje koji algoritam kodiranja se koristio za kodiranje sadržaja pre nego što je sadržaj poslat i prosleđuje se kroz Accept-Encoding polje zaglavlja. Svrha ovog procesa je kompresija čitavog sadržaja (dokumenta) koji se šalje, a da se pri tome ne izgube originalni podaci.
- Na tipove sadržaja odnosi se polje u zaglavlju koje nam ukazuje na to u kom formatu je sadržaj koji se šalje. Polje zaglavlja u kojem se šalju informacije o tipu sadržaja je Content-Type.
- Oznake jezika se šalju u Accept-Language ili Content-Language poljima zaglavlja.

